



A Testing Brief
*Setting Reading Standards,
Measuring Progress, and
Informing Instruction*

Leon R. Dreyfus, Ph.D.

Setting Reading Standards, Measuring Progress, and Informing Instruction

We have made great strides in understanding the critical elements required for children to become readers and we have begun to implement programs based on that understanding. Becoming a reader is not, however, the end of the story. With time, effort, and appropriate instruction, all of us can become better readers, and we expect our students to do so. With these good intentions and necessary resources, we are ready to act on the proposition that our students must read well in order to succeed. There are, however, some fairly basic questions to address. For example, what does it mean to read well? How do we know if students are doing it? And how do we know if they are moving in the right direction?

How well should our students be able to read?

There are really several parts to this question. What text should students be able to read and what should they be able to do with the text? Standards in many states address both parts of this question. With respect to the first part, most state standards mandate that: Students should be able to read and understand grade-level-appropriate material. But does that really tell us everything that we need to know? What are grade-level-appropriate materials? How well should students be able to read them? Does "grade-level-appropriate materials" cover everything that students should be able to read? To answer these questions fully, we need a way of characterizing the materials that students should be able to read and a way of characterizing the students' ability to read those materials.

Questar's **Degrees of Reading Power® (DRP) Program** does both. It characterizes the materials that students should be able to read in terms of the difficulty of the text they contain and characterizes student reading ability in terms of performance on a set of criterion-referenced tests.

This presentation will be organized in four parts to address the following issues:

- What should students be able to read and how do we set a standard in terms of those materials?
- How do we know whether or not the standard is being met?
- On the basis of the assessment information, what do we know about a student's reading ability?
- What are the instructional implications of this knowledge?

What should students be able to read and how do we set a standard in terms of those materials?

As a society, we have a certain set of expectations about what students should be able to read. Those expectations suggest a reformulation of the question as follows: what do students need to be able to read in order to be

A Testing Briefing
Setting Reading Standards, Measuring and Informing

successful in school, to be gainfully employed, to be literate citizens? We can specify those materials and use a readability formula to characterize them in terms of how difficult they are to read, or, in other words, their readability.

43 DRP Units	61 DRP Units	81 DRP Units
Many states are dry in summer. They get hardly any rain. Nearly all their water comes from melted snow. It is used during the growing season to water farms and orchards. Farmers buy the water. They are told how much they will be able to get. The amount changes each year. It depends on how snowy the winter was. A farmer needs to know how much he will receive. It allows him to decide which of several crops he ought to plant. The choice is based on how much water different crops need.	Natural gems are often treated to enrich their original colors. Stones such as jade and lapis lazuli are made darker by staining them. Over time, though, the color may fade. But fading is not unique to dyed gems. Some natural stones may lose color, too, if exposed to sunlight. Permanent changes in color can often be achieved by heating gems. A brown topaz, for example, turns a delicate pink when heated to 720 degrees Fahrenheit. Gemstones must be heated slowly, however, or they will crack.	Jefferson's preference for an agrarian society and his idealization of the independent farmer reflected a conviction that representative government required a secure and relatively prosperous economic base to function successfully. He perceived the farmer as economically independent, and thus unlikely to surrender his judgment as a citizen to the influence of demagogues. His dislike and distrust of cities derived from a conviction that urban conditions, especially for the poorer classes, forced men into such bitter struggle for sheer self-preservation that their natural moral sense could not be relied upon to produce social harmony or to guarantee responsible citizenship.

Continuous prose can be ordered in terms of its difficulty. The Degrees of Reading Power (DRP) Program reports the results of readability analyses in DRP units. Several short passages of text, presented above, illustrate differences in the readability of text.

There are several different types of materials that represent what we expect our students to be able to read. First, **students, as participating members of society, should be able to read the newspaper**. The analysis of text from newspapers around the country shows fairly consistent results. A sample of those results from different sections of several papers is shown below.

DRP® Difficulty of Articles Sampled from Several California Newspapers							
Newspapers	City	State	Year	Front Page	Editorial	Business	Sports
<i>San Jose Mercury News</i>	San Jose	CA	1998	69	69	72	60
<i>San Francisco</i>	San	CA	1998	70	67	71	60
<i>Los Angeles Times</i>	Los Angeles	CA	1998	71	71	70	62
<i>Sacramento Bee</i>	Sacramento	CA	1998	69	69	69	62

If they are going to drive a car, **students should also be able to read a driver's license manual.** We have analyzed the difficulty of the text in

Grade	Textbook Difficulty (Approximate range in DRP Units for 90% of books)
3	44-54
4	46-55
5	49-57
6	51-61
7	52-62
8	53-63
9	52-64
10	51-68
11	56-68
12	57-68

driver's license manuals from each of the 50 states. There is quite a bit of variability in the difficulty of the text in these manuals. The easiest (55 DRP units) would be at the independent reading level of the average 8th grader. The most difficult (70 DRP units) would be beyond the independent reading level of the average 12th grader.

At all educational levels, **students should be able to read their textbooks.** An extensive set of textbooks used in elementary, middle, high school, and college have been analyzed for text difficulty. Some of the results of those analyses are shown in the table at left.

Taken together, these analyses suggest that we do have a de facto expectation for reading proficiency. With this information, we can make explicit the expectations for student performance. For example, we might say that by the time students graduate from high school, they must be able to read independently the average college textbook. This amounts to a goal of 70 DRP units with 90% comprehension. Having set this standard, we can work backwards to determine where students should be at various points in their schooling. And, with the proper tools, we can monitor their progress towards this and intermediate goals.

How do we know whether or not the standard is being met?

Having formalized the expectations for what students should be able to read, we turn to the question of determining whether those standards have been met. The second element of the Degrees of Reading Power Program addresses this question with a series of criterion-referenced tests. In general terms, tests are designed to sort individuals. So, on the basis of a test we know how well one student performs relative to another. And with norm-referenced tests, we know how well a student performs relative to a large set of individuals. A criterion-referenced test, by contrast, measures student performance in terms of some meaningful educational outcome. DRP tests are criterion-referenced tests that measure a student's ability to comprehend the surface meaning of text. And that ability is reported on a scale of text difficulty, the same DRP scale that has been used to characterize the

A Testing Briefing
Setting Reading Standards, Measuring and Informing

difficulty of materials that students should be able to read. So DRP tests are the tools that we can use to determine whether the standard has been met.

DRP tests are holistic measures of reading comprehension that have the following characteristics. **DRP tests consist of a series of nonfiction paragraphs or passages.** Words have been deleted from the passages and the student is asked to select from a set of multiple-choice options the correct word for each deletion from the text. The passages and embedded items are carefully constructed so that the **paragraph or passage must be read and understood in order to answer correctly.** When a sentence with a blank is considered in isolation, each option is grammatically correct and semantically plausible. It is only in the context of the surrounding text that one option is unambiguously correct. In addition, **all of the content information that is needed to select the correct response is contained within the paragraph or passage.** The test measures the student’s knowledge of syntax, semantics, and other basic linguistic skills, not familiarity with the content of the passage. **DRP tests are untimed** so that slow, but accurate readers are not penalized. DRP tests report the most difficult text a student can read at a given level of comprehension. The DRP test was designed to measure progress, so **all forms measure the same construct and report results on the same scale.**

Grade	Textbook Difficulty (Range in DRP Units for 90% of books)	Time of Year	DRP® Reading Standards in Sample District	
			Remedial Standard	Reading Goal
3	44-54	Fall	29	37
		Spring	38	47
4	46-55	Fall	41	50
		Spring	45	54
5	49-57	Fall	46	55
		Spring	49	58
6	51-61	Fall	50	59
		Spring	53	62
7	52-62	Fall	54	63
		Spring	57	65
		Fall	58	67

The table at left shows how the DRP Program is used by one district to set standards and to monitor progress in reading. Several things are notable about this table. First, the standards make sense in terms of the difficulty of materials that students are expected to read in each grade. Second, the bar is raised each year to correspond to the

increasingly difficult materials that students will encounter. Third, this table

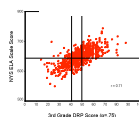
A Testing Briefing
Setting Reading Standards, Measuring and Informing

illustrates one set of standards. Obviously, other districts may establish different grade-level standards, depending on their curriculum and student population. In addition, other types of standards may be established. For example, other districts may choose to add an excellence standard. Finally, once these grade-level standards are established, whether an individual classroom, school, or district tests twice a year, once a year, or every two years, it is still possible to determine where each student stands with respect to these expectations. Of course, more frequent testing permits more timely intervention, but the information gained by more frequent testing must be balanced against the time and expense involved.

On the basis of information from DRP tests, what do we know about a student's reading ability? Having established 1) that there is a de facto set of expectations about reading that we can specify in terms of the text that students should be able to read, and 2) we can use the DRP readability analyses to formalize those standards and DRP tests to measure a student's progress, the next question is: What does the DRP test tell you about a student's reading ability? What it tells you is whether a student can use basic semantic, syntactic, and structural features of text to construct the meaning of grade-level-appropriate material. Most standards require more specific abilities, for example the abilities to analyze, evaluate, and criticize grade-level-appropriate text. The DRP test does not measure those latter abilities, but the knowledge of basic semantic, syntactic, and structural features of text that it does measure is fundamental to the successful completion of more advanced work with text.

If the DRP test does measure some fundamental reading ability, **a student's performance on the DRP test should show a strong, formal relation to tasks designed to measure abilities underlying more advanced work with text.** The figures below show the relation between performance on the DRP test and the New York State English Language Arts (ELA) test administered in 4th and 8th grades. ELA tests in both of these grades require, among other things, that students analyze, interpret, and critically evaluate grade-appropriate text.

**NYS 4th Grade ELA Test
Performance and 3rd Grade
DRP Scores**



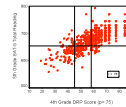
**NYS 8th Grade ELA Test
Performance and 8th Grade DRP
Scores**



The horizontal lines in both panels correspond to the boundary between performance levels 2 and 3 (out of 4) on the ELA test. The vertical lines were drawn to partition the data in order to facilitate predictions on the basis of DRP scores. These data show that there is a strong relationship between performance on the DRP test and ELA test and provide support for the contention that DRP tests measure a fundamental reading ability, one that is necessary for the performance of more advanced work with text.

Performance on the DRP test shows a strong relationship to other measures of academic success. The panel at the right shows how performance on the DRP test relates to the Total Reading Scale Score on the SAT9. This figure shows results for 550 students who took the DRP in the Spring of 4th grade and the SAT9 in 5th grade. Overall the correlation is higher than that between the DRP and either New York State test. The horizontal line was drawn at the 50th percentile on the SAT9. The vertical lines were drawn at DRP scores of 45 the 58. Using these criteria as minimum competency and excellence standards, respectively, would allow one to predict with fairly high accuracy whether a student was going to score above or below the 50th percentile on the SAT9. The point is not that these are the correct standards. Where one sets the standard depends on what sort of intervention is planned. The point is that what the DRP test measures is so fundamental that it can predict performance on a number of other instruments that measure more advanced abilities with text.

5th Grade SAT9 Total Reading and 4th Grade DRP Score (p=.75)



If the DRP test measures a fundamental ability then performance on the DRP test should be reflected in performance in the classroom. A study in Beacon, NY looked at the DRP scores received by 6th to 9th grade students and their final class grades. Students reading at the district “commencement” standard for the grade or higher (e.g., 70 DRP units, at p=.90, by the end of 9th grade)

received higher grades in all subjects – Math, Social Studies, Science, and English – than students reading below this standard.

What are the instructional implications of the knowledge provided by DRP tests?

We know that students learn the most and grow the most as readers when reading material provides the proper blend of success and challenge. When

text
is
too

Grade	Textbook Difficulty Average in DRP Units	Reading Ability (DRP Scores) at each Quartile for Fall of the Academic Year					
		Instructional (P = .75)			Independent (P = .90)		
		25 th	50 th	75 th	25 th	50 th	75 th
6	56	47	56	66	36	45	55
7		51	61	72	40	50	61
8		55	65	75	44	54	64
9	62	58	68	78	47	57	67
10		60	70	80	49	59	69
11		62	72	82	51	61	71

difficult, students will become frustrated and the motivation to read will be undermined. When text is too easy, students are not exposed to more advanced vocabulary and syntactic structures. With DRP scores, teachers can know how a student's ability compares to the difficulty of textbooks and other materials that students are asked to read. Recently, Allington (2002) has suggested that many students in grades 5 through 12 struggle to learn from content-area textbooks that are too difficult for them. The DRP program, because it reports reading ability on a scale of text difficulty, can be used to evaluate the extent of the problem. The table above shows, for grades 6 through 12, a summary of the most recent DRP test norms and permits a comparison of student ability with the difficulty of the average middle school and high school textbooks. This analysis shows that the reading ability of a substantial number of students is well below the difficulty of the textbook that they are expected to read. For example, **in the 9th grade the independent reading ability of the average student, DRP (p=.90) = 57, is lower than the difficulty of the average high school textbook, i.e., 62 DRP units. For students whose ability is below average, this discrepancy is even larger.**

With the knowledge provided by the DRP program, teachers can decide where and what form of extra assistance may be needed. When the text is within range of the student's instructional reading level, strategies focusing on vocabulary and text structures of the text may be sufficient. When the gap between ability and text difficulty is larger, more extreme measures may be required. Allington (2002) makes a strong case for a multi-sourced and multi-leveled curriculum. **DRP->BookLink**, another element of the DRP program, can be used as a resource in building such curricula. It allows teachers to select books in general as well as specific topic areas; but, more to the point, they can select books that are matched to each student's ability as well as to their interests.

The DRP program is not prescriptive with respect to the instructional approach used to solve the problem of textbooks that are not well matched to the abilities of students. It can be used, however, to identify and to determine the extent of the problem and to measure the success of intervention programs.

Reference

Allington, R. L. (2002). You can't learn much from books you can't read. *Educational Leadership*, 60 (3), 16-19.